

CROATICA CHEMICA ACTA
CCACAA **80** (2) 217–225 (2007)

ISSN-0011-1643

CCA-3163

Original Scientific Paper

Numerical Representation of Three-dimensional Stereochemical Environments Using FRAU-descriptors*

Hiroko Satoh

*National Institute of Informatics, 2-1-2 Hitotsubashi, Chiyoda-ku, Tokyo 101-8430, Japan
(E-mail: hsatoh@nii.ac.jp)*

RECEIVED OCTOBER 5, 2006; REVISED MARCH 8, 2007; ACCEPTED MARCH 14, 2007

<p><i>Keywords</i> FRAU-descriptors 3D stereochemical environments quantification axial and equatorial orientations</p>	<p>Numerical representation of stereochemical environments using FRAU (Field-characterization for Reaction Analysis and Understanding)-descriptors has been investigated. Correlations were analyzed between the FRAU-descriptors and stereochemical environments around an oxygen atom of cyclohexanol with chair form and oxygen atoms of all of the diastereo isomers of 1,2-cyclohexanediol, 1,3-cyclohexanediol, 1,4-cyclohexanediol, and 1,2,3,4,5-cyclohexanepentol with all possible chair form conformations. Good correlations were found between the FRAU-descriptors and 3D stereochemical environments.</p>
---	--

INTRODUCTION

Stereochemistry is essential information for discussing property, activity and reactivity based on chemical structures. Many research groups have developed methods for handling stereochemical information using computers,^{1–12} these methods have been for linear notations of stereochemical structures, character-labels representing stereochemical attributes, and numeric representation of stereochemical environments.

My colleagues and I developed two methods for stereochemical representation called FRAU (Field-characterization for Reaction Analysis and Understanding)⁹ and CAST (CAnonical-representation for STereochemistry).¹⁰ The CAST method provides canonical line notations representing stereochemical information and was implemented in a ¹³C NMR chemical shift prediction system called CAST/CNMR,¹³ which ensures highly accurate prediction by considering stereochemistry, this system has been

used in practical structural analysis such as structural revision.¹⁴ The FRAU system numerically characterizes molecular fields based on electrostatic and steric interactions and was applied to classifying and predicting reagents' functions.¹⁵

To determine the further applicability of FRAU for detecting more detailed similarities and differences in reagents and also for predicting stereoselectivity and regioselectivity, I have investigated whether the 3D stereochemical environments can be distinguished with the FRAU-descriptors.

EXPERIMENTAL

FRAU-descriptors

The FRAU system provides three kinds of numerical descriptors concerning electrostatic interactions, steric interactions, and surface area, which are called FFElectro, FFsteric, and

* Dedicated to Professor Haruo Hosoya in happy celebration of his 70th birthday.

FFfield, respectively. These descriptors characterize a reaction field in the three-dimensional space and are assigned to each atom in a molecule. The FFelectro descriptor is calculated based on electrostatic interactions, the FFsteric descriptor is calculated based on van der Waals interactions, and the FFfield descriptor corresponds to the size of the van der Waals surface. The size of the area is estimated as the number of dots that are equally set on the surface. A detailed FRAU algorithm was described in a previous paper.⁹

The purpose of this study is to determine whether the FFsteric and FFfield descriptors can detect the differences and similarities of stereochemical environments.

Dataset

The model structures used in this study are shown in Figures 1–5. Correlations were analyzed between the two types of FRAU-descriptors and stereochemical environments around oxygen atoms of cyclohexanol (**1**) (Figure 1) and of all diastereo isomers of 1,2-cyclohexanediol (**2**) (Figure 2), 1,3-cyclohexanediol (**5**) (Figure 3), 1,4-cyclohexanediol (**8**) (Figure 4), and 1,2,3,4,5-cyclohexanepentol (**11**) (Figure 5) with chair form conformations. These model structures were chosen because they contain oxygen atoms with various 3D stereochemical environments that form an appropriate dataset to investigate the correlation. In total, the dataset consisted

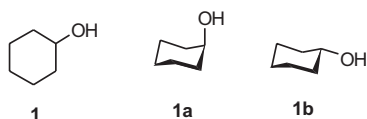


Figure 1. Structures of cyclohexanol (**1**) and its chair form conformers (**1a**, **1b**).

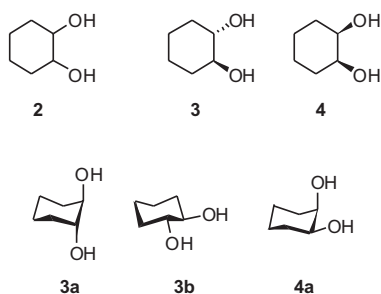


Figure 2. Structures of 1,2-cyclohexanediol (**2**), its diastereo isomers (**3**, **4**), and their chair form conformers (**3a**, **3b**, **4a**).

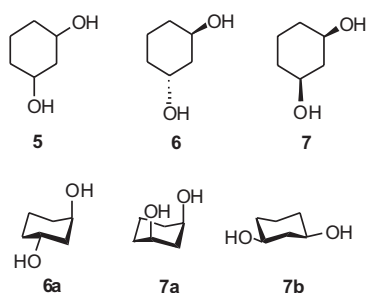


Figure 3. Structures of 1,3-cyclohexanediol (**5**), its diastereo isomers (**6**, **7**), and their chair form conformers (**6a**, **7a**, **7b**).

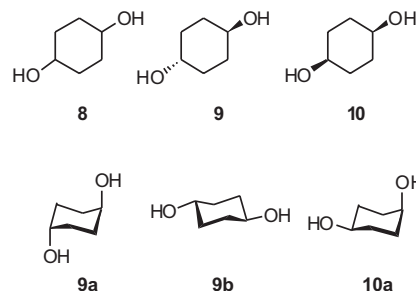


Figure 4. Structures of 1,4-cyclohexanediol (**8**), its diastereo isomers (**9**, **10**), and their chair form conformers (**9a**, **9b**, **10a**).

of FFsteric and FFfield descriptors of 120 oxygen atoms. In calculation of the descriptors, in order to ignore differences in the orientations of hydrogen atoms, no hydrogen atoms of the hydroxyl groups were taken into account.

RESULTS AND DISCUSSION

Plot Graph of All Data

The plot graph of FFsteric to FFfield of all of the 120 oxygen atoms is shown in Figure 6-a. Equatorial and axial atoms were clearly classified, and the axial atoms were grouped into axial, 1,3-diaxial, and 1,3,5-diaxial classes, as shown in Figure 6-b. We looked into the detailed contents of the classes as follows.

Contents of the Clusters of Equatorial Atoms

The data corresponding to the Ea–f clusters in the equatorial group (Figure 6-b) are shown in Figures 7–12. The Ea cluster consisted of four equatorial oxygen atoms of **1b**, **6a**, **9b**, and **10a**, as shown in Figure 7, where corresponding atoms are denoted by gray circles. The common partial structure around the equatorial atoms is shown in Figure 7-a, they have no hydroxyl group at either of the β -carbon atoms. The Eb cluster consisted of 11 equatorial oxygen atoms that are denoted by gray circles in the **4a**, **12a**, **13a**, **14a**, **15a**, **16a**, **16b**, **17a**, and **20a** structures (Figure 8). The characteristic partial structure for the cluster had an axial hydroxyl group at one of the β -carbon atoms (Figure 8-a). The Ec cluster was characterized by a partial structure that had an equatorial hydroxyl group at one of the β -carbon atoms, as shown in Figure 9-a; that partial structure was the common structure around 11 of the equatorial atoms denoted by gray circles in **3b**, **13b**, **14a**, **17b**, **18a**, **18b**, **19b**, **20a**, and **21a** (Figure 9). The Ed cluster consisted of 8 equatorial oxygen atoms that had axial hydroxyl groups at both of the β -carbon atoms (Figure 10-a), which are in **12a**, **12b**, **13a**, **13b**, **14b**, **16b**, and **19a** (Figure 10). The Ee cluster was characterized by a partial structure that had an equatorial hydroxyl group at one of the β -carbon atoms and an axial hydroxyl group at the other β -carbon atom, as shown in Figure 11-a; that was the common structure around 14 of

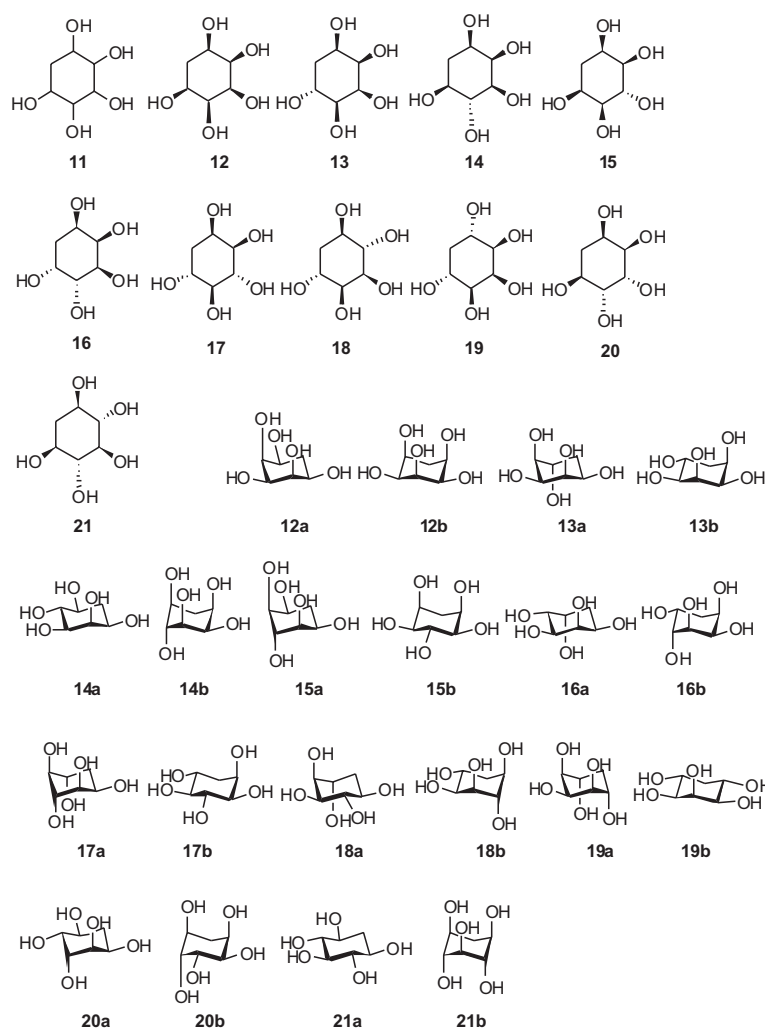


Figure 5. Structures of 1,2,3,4,5-cyclohexanepentol (**11**), its diastereo isomers (**12–20**) and their chair form conformers (**12a**, **12b**, **13a**, **13b**, **14a**, **14b**, **15a**, **15b**, **16a**, **16b**, **17a**, **17b**, **18a**, **18b**, **19a**, **19b**, **20a**, **20b**, **21a**, **21b**).

the equatorial atoms denoted by gray circles in **13b**, **14a**, **15b**, **16a**, **17b**, **18a**, **18b**, **19b**, **20a**, and **20b** (Figure 11). The Ef cluster consisted of eight equatorial oxygen atoms that had the same partial structure with equatorial hydroxyl groups at both of the β -carbon atoms (Figure 12-a), which are in **14a**, **15b**, **17b**, **18a**, and **21a**.

Figure 13 shows the distance from the equatorial atoms in the Ea cluster to the other atoms. This shows that the equatorial oxygen atoms share the same three-dimensional partial structural environment within about 4×10^{-10} m, which corresponds to the partial structure characterizing the Ea cluster shown in Figure 7-a. For the other clusters, the corresponding equatorial oxygen atoms also share the same three-dimensional partial structural environment within about 4×10^{-10} m.

The results show that the FFsteric and FFfield descriptors clearly detected the similarities of the three-dimensional partial structural environments around the oxygen atoms; the size of the environments was within about 4×10^{-10} m.

Contents of the Clusters of Axial Atoms

In the same way shown for the equatorial atoms, the data corresponding to the Aa–c, 13Aa–c, and 135Aa–c clusters in the axial group, the 1,3-diaxial group, and 1,3,5-diaxial group (Figure 6-b) are shown in Figures 14–22, respectively. The Aa, Ab, and Ac clusters were characterized by partial structures that had no equatorial hydroxyl groups at either of the β -carbon atoms (Figure 14-a), those that had an equatorial hydroxyl group at one of the β -carbon atoms (Figure 15-a), and those that had equatorial hydroxyl groups at both of the β -carbon atoms (Figure 16-a). The 13Aa, 13Ab, and 13Ac clusters were characterized by partial structures that had no equatorial hydroxyl groups at either of the β -carbon atoms (Figure 17-a), those that had an equatorial hydroxyl group at one of the β -carbon atoms (Figure 18-a), and those that had equatorial hydroxyl groups at both of the β -carbon atoms (Figure 19-a). The 135Aa, 135Ab, and 135Ac clusters were characterized by partial structures that had no equatorial hydroxyl groups at either of the β -carbon atoms (Figure 20-a), those

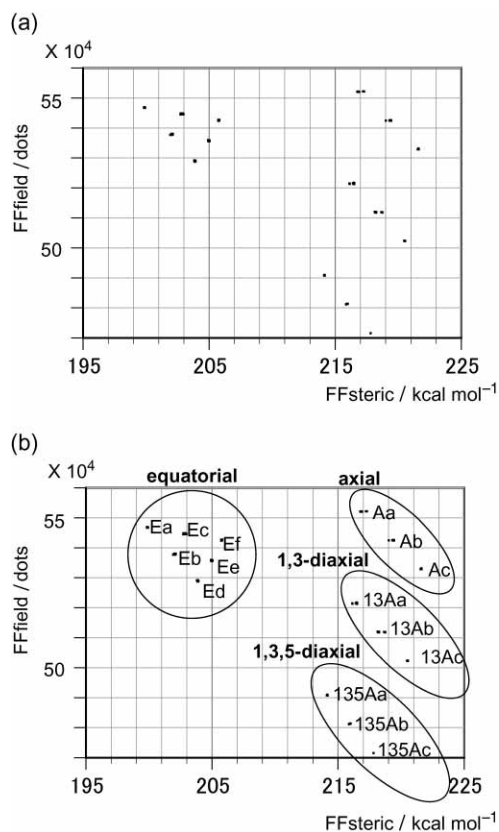


Figure 6. Plot graph of FFsteric to FFfield for all 120 oxygen atoms. (a) A plot graph. (b) The oxygen atoms were grouped into equatorial, axial, 1,3-diaxial and 1,3,5-diaxial classes.

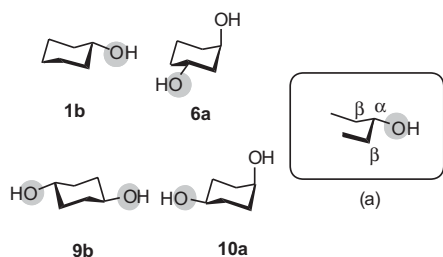


Figure 7. Contents of the Ea cluster. (a) Characteristic partial structure.

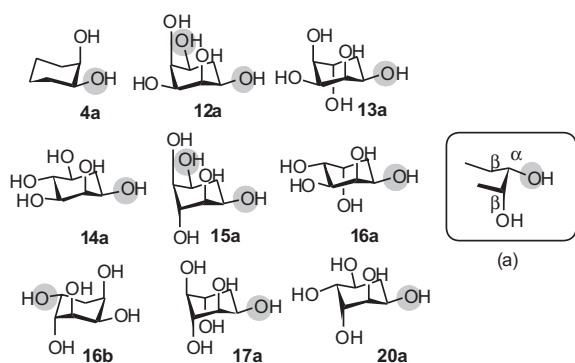


Figure 8. Contents of the Eb cluster. (a) Characteristic partial structure.

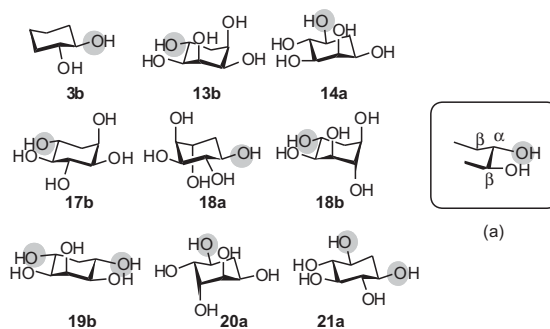


Figure 9. Contents of the Ec cluster. (a) Characteristic partial structure.

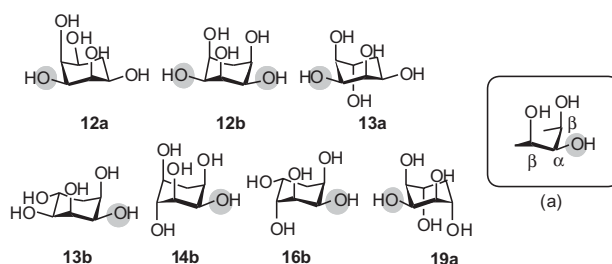


Figure 10. Contents of the Ed cluster. (a) Characteristic partial structure.

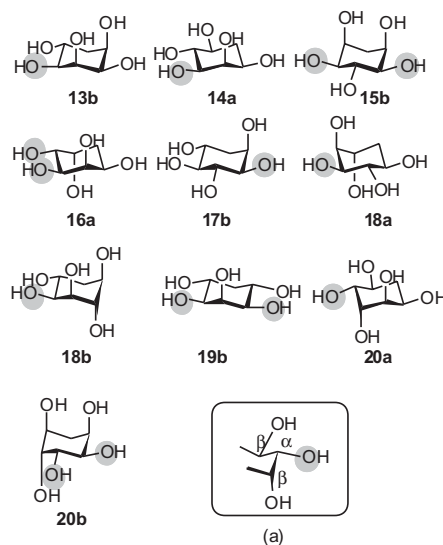


Figure 11. Contents of the Ee cluster. (a) Characteristic partial structure.

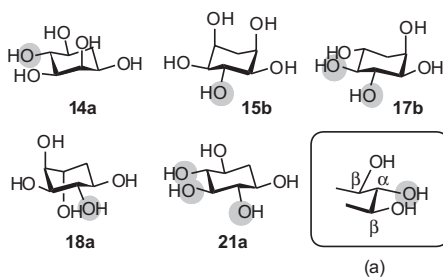


Figure 12. Contents of the Ef cluster. (a) Characteristic partial structure.

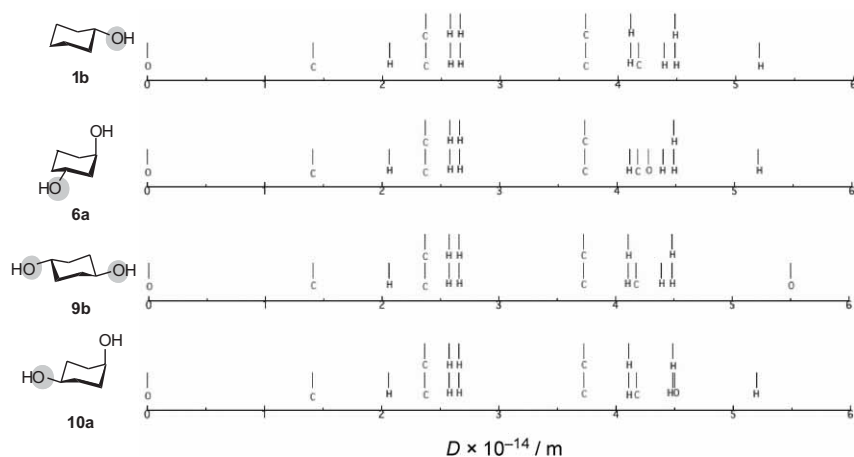


Figure 13. Distance from the equatorial atoms in the Ea cluster to the other atoms.

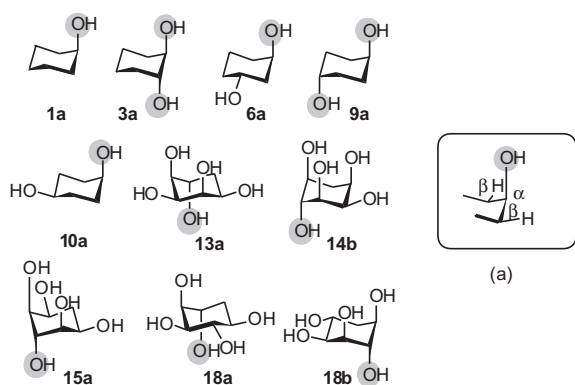


Figure 14. Contents of the Aa cluster. (a) Characteristic partial structure.

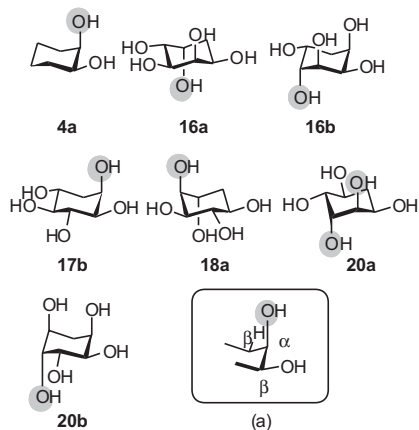


Figure 15. Contents of the Ab cluster. (a) Characteristic partial structure.

that had an equatorial hydroxyl group at one of the β -carbon atoms (Figure 21-a), and those that had equatorial hydroxyl groups at both of the β -carbon atoms (Figure 22-a).

Thus, the An , $13An$, and $135An$ clusters ($n = a, b$, and c) are characterized by the same partial structure but have

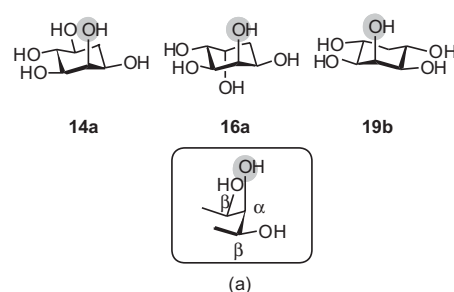


Figure 16. Contents of the Ac cluster. (a) Characteristic partial structure.

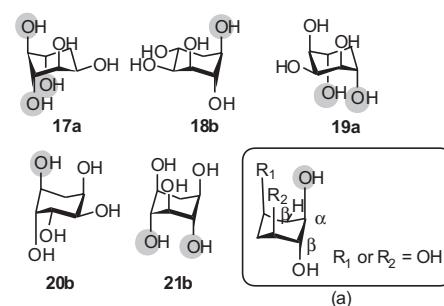


Figure 17. Contents of the 13Aa cluster. (a) Characteristic partial structure.

different numbers of diaxial oxygen atoms in their three-dimensional space. The results show that the FFsteric and FFfield descriptors also clearly detected the similarities of three-dimensional structural environments for the axial oxygen atoms.

Correlations among the Clusters

I found some interesting linear relationships among the clusters by looking into the plot graph (Figure 6).

The equatorial group includes three types of linear relationships (Figures 23, 24, and 25). Figure 23 shows the linear relationship of Ea, Ec, and Ef (line (a)), and that of Eb and Ee (line (b)). The characteristic partial

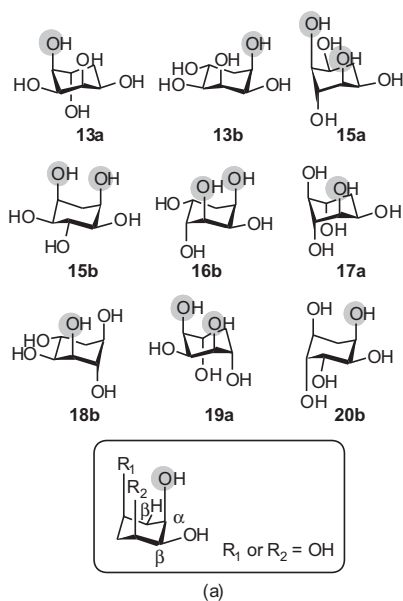


Figure 18. Contents of the 13Ab cluster. (a) Characteristic partial structure.

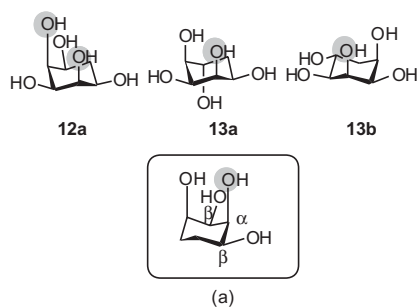


Figure 19. Contents of the 13Ac cluster. (a) Characteristic partial structure.

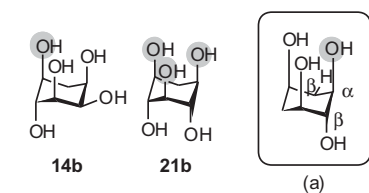


Figure 20. Contents of the 135Aa cluster. (a) Characteristic partial structure.

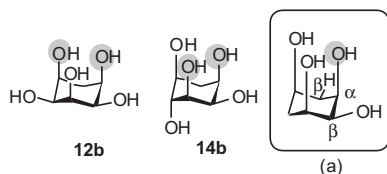


Figure 21. Contents of the 135Ab cluster. (a) Characteristic partial structure.

structures corresponding to lines (a) and (b) are shown in Figure 23-i and ii, respectively. These show that the number of equatorial carbon atoms at the β -carbon atoms

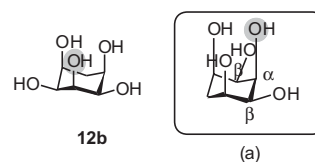


Figure 22. Contents of the 135Ac cluster. (a) Characteristic partial structure.

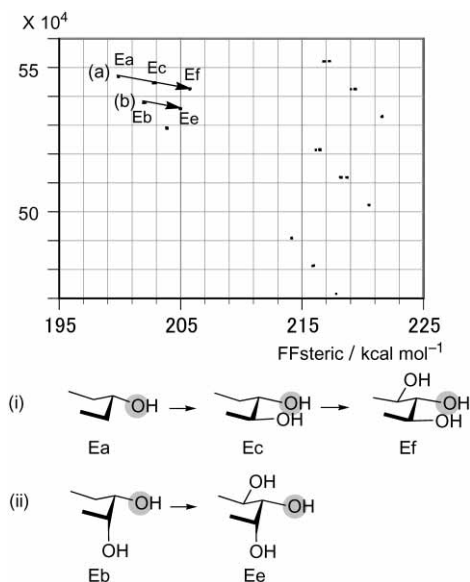


Figure 23. Equatorial addition effect: Linear relationship in the equatorial group when the number of equatorial hydroxyl groups at the β -carbon atoms increases.

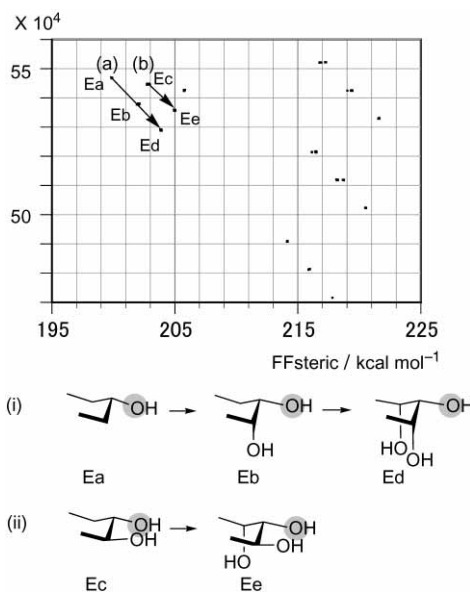


Figure 24. Axial addition effect: Linear relationship in the equatorial group when the number of axial hydroxyl groups at the β -carbon atoms increases.

increases one by one from Ea to Ef in line (a) and from Eb to Ee in line (b). The lines share the same gradient. The results show that the FFsteric and FFfield descriptors

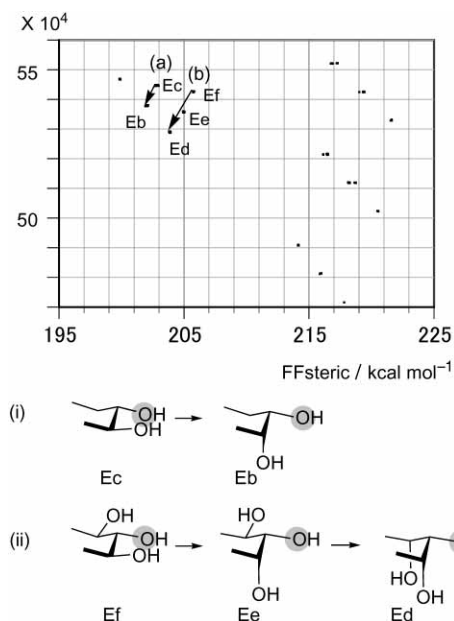


Figure 25. Equatorial to axial change effect: Linear relationship in the equatorial group when the orientation of a hydroxyl group at the β -carbon atom changes from equatorial to axial.

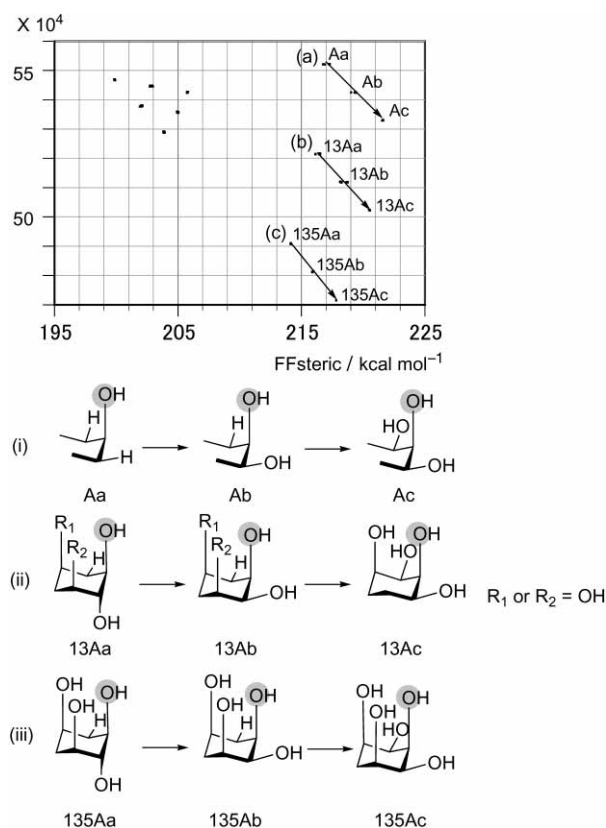


Figure 26. Equatorial addition effect: Linear relationship in the axial group when the number of equatorial hydroxyl groups at the β -carbon atoms increases.

linearly detected the increase in the number of equatorial hydroxyl groups at β -carbon atoms.

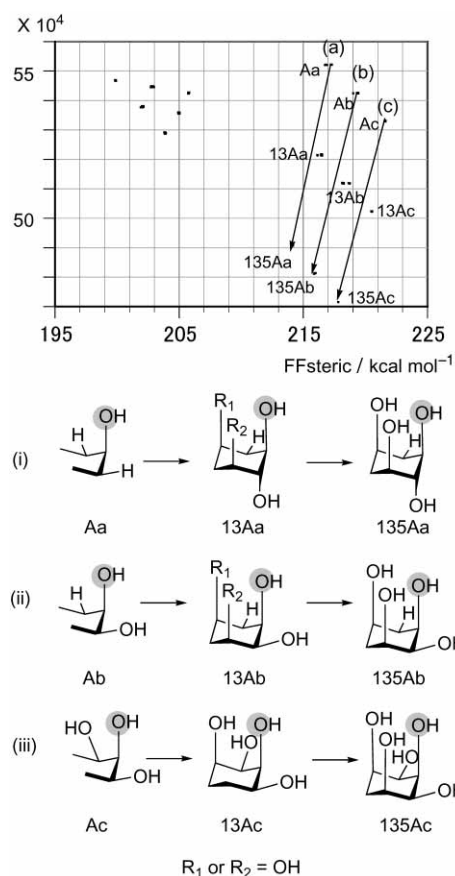


Figure 27. Diaxial addition effect: Linear relationships in the axial group when the number of diaxial hydroxyl groups at the β -carbon atoms increases.

Figure 24 shows the linear relationship of E_a , E_b , and E_d (line (a)), and that of E_c and E_e (line (b)), of which characteristic partial structures are shown in Figure 24-i and ii, respectively. They show that the number of axial carbon atoms at the β -carbon atoms increases one by one from E_a to E_d , and from E_c to E_e . The lines share the same gradient. The results show that the FF_{steric} and FF_{field} descriptors also linearly detected the increase in the number of axial hydroxyl groups at the β -carbon atoms.

Figure 25 shows the linear relationship of E_c and E_b (line (a)) and that of E_f , E_e , and E_d (line (b)). The characteristic partial structures corresponding to lines (a) and (b) shown in Figure 25-i and ii, respectively. These show that a hydroxyl group at a β -carbon atom changes from an equatorial orientation into axial orientation one by one from E_c to E_b , and from E_f to E_d . The lines share the same gradient. The results show that the FF_{steric} and FF_{field} descriptors linearly detected the changes in the orientation of a hydroxyl group at a β -carbon atom.

The axial group includes two types of linear relationships (Figures 26 and 27). Figure 26 shows the linear relationship of A_a , A_b , and A_c (line (a) in Figure 26), that of $13A_a$, $13A_b$, and $13A_c$ (line (b) in Figure 26), and that

of 135Aa, 135Ab, and 135Ac (line (c) in Figure 26). The characteristic partial structures corresponding to lines (a) and (b) are shown in i and ii, respectively. These show that the number of equatorial hydroxyl groups at the β -carbon atoms increases one by one from Aa to Ac, from 13Aa to 13Ac, and from 135Aa to 135Ac. The lines share the same gradient, which is the same as the gradient of the lines in Figure 24, this consistent gradient demonstrates the relationships of the increase in the number of axial hydroxyl groups at the β -carbon atoms for equatorial oxygen atoms. The results show that the FFsteric and FFfield descriptors linearly detected the increase in the number of equatorial hydroxyl groups at the β -carbon atoms. The descriptors also numerically represented the relationships between axial and equatorial hydroxyl groups in the three dimensional space.

The other type of linear relationships in axial group is demonstrated by Aa, 13Aa, and 135Aa (line (a) in Figure 27), of Ab, 13Ab, and 135Ab (line (b) in Figure 27), and of 135Aa, 135Ab, and 135Ac (line (c) in Figure 27). The characteristic partial structures corresponding to lines (a), (b), and (c) are shown in i, ii, and iii, respectively. These show that the number of diaxial hydroxyl groups increases one by one from Aa, 13Aa, and 135Aa, from Ab, 13Ab, and 135Ab, and from Ac, 13Ac, and 135Ac. The gradients of the lines are about the same. The results show that the FFsteric and FFfield descriptors linearly detected the differences in the number of the diaxial hydroxyl group. When the number of diaxial groups increase, both FFfield and FFsteric decrease. The decrease in FFfield, which is the number of dots on the atomic surface, is reasonable but the decrease in FFsteric, which is van der Waals intermolecular interaction energy, is not reasonable because if the number of diaxial groups is increased, the steric repulsive energy will be high. The reason for the decrease in FFsteric may be that the interaction in the space near the α -carbon atom can not be detected in the case of the 1,3- and 1,3,5-diaxial structures. This is because the van der Waals surface is not near the α -carbon atom due to the highly steric hindrance. This is a problem in treating highly hindered structures and will be improved in the future version of the FRAU system.

CONCLUSION

Correlations were investigated between the steric and field FRAU (Field-characterization for Reaction Analysis and Understanding)-descriptors and stereochemical environments around oxygen atoms in chair form conformations

of cyclohexanol and all diastereo isomers of 1,2-cyclohexanediol, 1,3-cyclohexanediol, 1,4-cyclohexanediol, and 1,2,3,4,5-cyclohexanepentol. Detailed partial conformational similarities and differences were detected using the FRAU-descriptors. The results demonstrated the FRAU system's high potential to be applied to quantify factors concerning conformational and 3D structural features. These functions will be useful for several structure-based prediction models concerning reaction prediction, NMR estimation, and quantitative structure-activity relationships analyses.

Supporting Information Available – All of the calculated FFsteric and FFfield descriptors (4 pages) are available on request from cheminfo@nii.ac.jp.

REFERENCES

1. W. T. Wipke and T. M. Dyott, *J. Am. Chem. Soc.* **96** (1974) 4834–4842.
2. F. Choplin, R. Dorschner, G. Kaufmann, and W. T. Wipke, *J. Organometallic Chem.* **152** (1978) 101–109.
3. W. C. Herndon and J. E. Leonard, *Inorg. Chem.* **22** (1983) 554–557.
4. T. Akutsu, *J. Chem. Inf. Comput. Sci.* **31** (1991) 414–417.
5. M. Razinger and M. Perdih, *J. Chem. Inf. Comput. Sci.* **34** (1994) 290–296.
6. K. K. Agarwal, *J. Chem. Inf. Comput. Sci.* **34** (1994) 463–479.
7. P. Mata and A. M. Lobo, *J. Chem. Inf. Comput. Sci.* **34** (1994) 491–504.
8. J. L. Markley, A. Bax, Y. Arata, C. W. Hilbers, R. Kaptein, B. D. Sykes, P. E. Wright, and K. Wüthrich, *J. Biomol. NMR* **12** (1998) 1–23.
9. H. Satoh, S. Itono, K. Funatsu, K. Takano, and T. Nakata, *J. Chem. Inf. Comput. Sci.* **39** (1999) 671–678.
10. (a) H. Satoh, H. Koshino, K. Funatsu, and T. Nakata, *J. Chem. Inf. Comput. Sci.* **40** (2000) 622–630; (b) H. Satoh, H. Koshino, K. Funatsu, and T. Nakata, *J. Chem. Inf. Comput. Sci.* **41** (2001) 1106–1112; (c) H. Satoh, H. Koshino, and T. Nakata, *J. Comput. Aided Chem.* **3** (2002) 48–55.
11. S. Fujita, *Computer-Oriented Representation of Organic Reactions*, Yoshioka Shoten Publishing Company, 2001.
12. J. Aires-de-Sousa and J. Gasteiger, *J. Mol. Graphics Modell.* **20** (2002) 373–388.
13. (a) H. Satoh, H. Koshino, J. Uzawa, and T. Nakata, *Tetrahedron* **59** (2003) 4539–4547; (b) H. Satoh, H. Koshino, T. Uno, S. Koichi, S. Iwata, and T. Nakata, *Tetrahedron* **61** (2005) 7431–7437.
14. H. Koshino, H. Satoh, T. Yamada, and Y. Esumi, *Tetrahedron Letters* **47** (2006) 4623–4626.
15. H. Satoh, K. Funatsu, K. Takano, and T. Nakata, *Bull. Chem. Soc. Jpn.* **73** (2000) 1955–1965.

SAŽETAK

Numerički prikaz trodimenzionalnih stereokemijskih okolina korištenjem FRAU opisivača

Hiroko Satoh

U radu se istražuje numerički prikaz stereokemijskih okolina pomoću tzv. FRAU opisivača sa stereokemijskim okolinama kisikovog atoma u cikloheksanolu i njemu pripadnim diastereoizomerima za sve njihove moguće konformacije stolice. Utvrđena je dobra koleracija između FRAU opisivača i 3D-stereokemijskih okolina.

Supporting Information

Numerical Representation of Three-dimensional Stereochemical Environments Using FRAU-descriptors

Hiroko Satoh

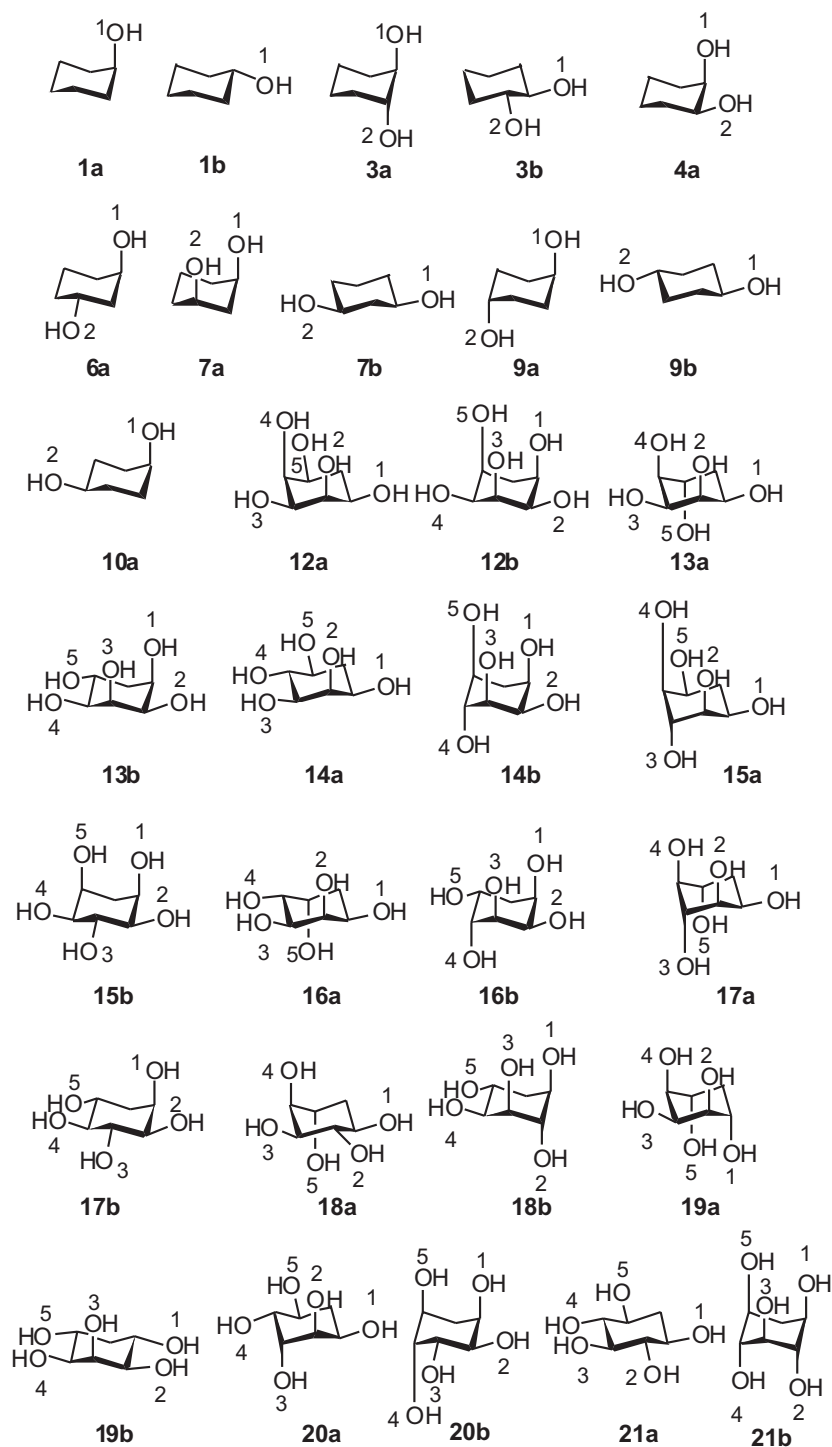
National Institute of Informatics, 2-1-2 Hitotsubashi, Chiyoda-ku, Tokyo 101-8430, Japan

Figure S1. The model structures with serial nos. of oxygen atoms.

TABLE SI. FFsteric and FFfield descriptors for all of the structures

Compound no.	Atom no.	FFsteric kcal mol ⁻¹	FFfield	Compound no.	Atom no.	FFsteric kcal mol ⁻¹	FFfield
1a	1	216.68491	552688	14b	1	215.80222	481860
1b	1	199.77820	547568		2	203.82912	529759
3a	1	216.82715	552756		3	215.9239	481977
	2	216.82552	552654		4	217.14496	552756
3b	1	216.82706	552756		5	214.10377	491362
	2	216.82542	552654	15a	1	202.03111	538657
4a	1	202.6763	545449		2	218.66907	512526
	2	202.65922	545107		3	217.21591	552756
6a	1	218.95812	543089		4	218.67643	512388
	2	201.85067	538286		5	202.05454	538456
7a	1	216.80783	552654	15b	1	218.6632	512488
	2	199.85153	547279		2	204.91345	536559
7b	1	216.0784	521965		3	205.73839	543323
	2	216.06404	522103		4	204.92236	536176
9a	1	199.82714	547279		5	218.67125	512401
	2	199.8248	547568	16a	1	201.96634	538657
9b	1	216.73691	552654		2	221.48168	533512
	2	216.76243	552759		3	204.90688	536161
10a	1	199.79129	547568		4	204.90882	536212
	2	199.77538	547376		5	219.28451	543200
12a	1	202.00471	538657	16b	1	218.06717	512653
	2	220.46729	502961		2	203.81188	529803
	3	203.77938	529355		3	218.19733	512570
	4	220.47323	502807		4	219.30503	543004
	5	202.02814	538456		5	201.99785	538286
12b	1	215.88832	482012	17a	1	202.05598	538657
	2	203.82155	529612		2	218.60883	512526
	3	217.74515	472145		3	216.45477	522180
	4	203.85875	529475		4	216.45467	521965
	5	215.81193	481761		5	216.36844	522131
13a	1	202.02957	538657	17b	1	219.34432	543049
	2	220.40583	502961		2	204.8502	536559
	3	203.80622	529355		3	205.71469	543323
	4	218.17907	512384		4	205.70297	542982
	5	217.15138	552759		5	202.80878	545149
13b	1	218.13026	512420	18a	1	202.84284	545457
	2	203.74517	529759		2	205.75096	543323
	3	220.48085	502963		3	204.91631	536176
	4	204.91316	536161		4	219.37501	543003
	5	202.8326	545149		5	217.22678	552759
14a	1	201.94148	538657	18b	1	216.39924	521998
	2	221.55986	533512		2	217.20528	552797
	3	204.88263	536161		3	218.67833	512540
	4	205.67905	543024		4	204.93941	536161
	5	202.83641	545257		5	202.90194	545149

Compound no.	Atom no.	FFsteric kcal mol ⁻¹	FFfield
19a	1	216.34774	522220
	2	218.10575	512538
	3	203.83307	529355
	4	218.11877	512384
	5	216.38369	522148
19b	1	202.80816	545397
	2	204.86277	536570
	3	221.63255	533747
	4	204.8887	536396
	5	202.82749	545202
20a	1	201.96788	538657
	2	219.35706	543077
	3	219.35192	543206
	4	204.9087	536232
	5	202.86024	545257
20b	1	218.58774	512536
	2	204.95526	536413
	3	204.96823	536319
	4	219.24928	543095
	5	216.39494	521902
21a	1	202.74931	545457
	2	205.6579	543323
	3	205.66666	542982
	4	205.65537	543024
	5	202.76713	545257
21b	1	214.08337	491605
	2	216.41859	522328
	3	214.10074	491429
	4	216.37843	522102
	5	214.05316	491398